DEQOR: a web-based tool for the design and quality control of siRNAs

Andreas Henschel, Frank Buchholz[1] and Bianca Habermann*

Scionics Computer Innovation, GmbH, Pfotenhauerstrasse 110, 01307 Dresden, Germany and [1]Max Planck Institute of Molecular Cell Biology and Genetics, Pfotenhauerstrasse 108, 01307 Dresden, Germany

ABSTRACT

RNA interference (RNAi) is a powerful tool for inhibiting the expression of a gene by mediating the degradation of the corresponding mRNA. The basis of this gene-specific inhibition is small, double-stranded RNAs (dsRNAs), also referred to as small interfering RNAs (siRNAs), that correspond in sequence to a part of the exon sequence of a silenced gene. The selection of siRNAs for a target gene is a crucial step in siRNA-mediated gene silencing. According to present knowledge, siRNAs must fulfill certain properties including sequence length, GC-content and nucleotide composition. Furthermore, the cross-silencing capability of dsRNAs for other genes must be evaluated. When designing siRNAs for chemical synthesis, most of these criteria are achievable by simple sequence analysis of target mRNAs, and the specificity can be evaluated by a single BLAST search against the transcriptome of the studied organism. A different method for raising siRNAs has, however, emerged which uses enzymatic digestion to hydrolyze long pieces of dsRNA into shorter molecules. These endoribonuclease-prepared siRNAs (esiRNAs or 'diced' RNAs) are less variable in their silencing capabilities and circumvent the laborious process of sequence selection for RNAi due to a broader range of products. Though powerful, this method might be more susceptible to cross-silencing genes other than the target itself. We have developed a web-based tool that facilitates the design and quality control of siRNAs for RNAi. The program, DEQOR, uses a scoring system based on state-of-the-art parameters for siRNA design to evaluate the inhibitory potency of siRNAs. DEQOR, therefore, can help to predict (i) regions in a gene that show high silencing capacity based on the base pair composition and (ii) siRNAs with high silencing potential for chemical synthesis. In addition, each siRNA arising from the input query is evaluated for possible cross-silencing activities by performing BLAST searches against the transcriptome or genome of a selected organism. DEQOR can therefore predict the probability that an mRNA fragment will cross-react with other genes in the cell and helps researchers to design experiments to test the specificity of esiRNAs or chemically designed siRNAs. DEQOR is freely available at http://cluster-1.mpi-cbg.de/Deqor/deqor.html.

INTRODUCTION

RNA interference (RNAi) has emerged as a powerful tool for the sequence-specific silencing of mRNAs in eukaryotic cells [reviewed in (1,2)]. First discovered in plants and nematodes, it soon became clear that RNAi is conserved throughout the eukaryotic kingdom as a means of gene regulation and protection of an organism against parasites such as viruses and transposons [reviewed in (3–7)]. The machinery involved in RNAi consists of a complex mixture that includes a helicase, endo- and exonucleases and most likely—at least in the nematode *Caenorhabditis elegans*—an RNA-directed RNA polymerase [reviewed in (2)]. The endoribonuclease Dicer triggers RNAi response by the digestion of long double-stranded RNA (dsRNA) into small pieces of ∼21 nt length that are referred to as small interfering RNAs (siRNAs). Invertebrates such as *C.elegans* and *Drosophila melanogaster* are able to effectively digest and therefore employ long pieces of dsRNA for dsRNA-based gene silencing. Several genome-wide RNAi screens have therefore been carried out successfully in these invertebrate model systems for the discovery of new genes in different cellular pathways (8–15). The introduction of long dsRNA into mammalian systems is, however, more problematic. This is due to the fact that in most mammalian cells dsRNA that is longer than 40 nt (Frank Buchholz, unpublished data) induces a nonspecific interferon response, leading to the general

*To whom correspondence should be addressed. Tel: +49 351 210 2543; Fax: +49 351 210 2000; Email: habermann@mpi-cbg.de

shutdown of transcription and/or cell death (16,17). Most standard protocols for RNAi in mammalian systems therefore use chemically synthesized siRNAs, and this method has emerged as a promising tool for sequence-specific gene silencing in mammalian cell culture (18). Yet this approach is limited by the fact that different sequences within a gene have dramatically varied inhibitory abilities (19). In essence, a large number of different synthetic siRNAs have to be screened for their efficacy at knocking down the gene of interest, which is a laborious and costly task. The applicability for high-throughput screens using chemically synthesized siRNAs is therefore questionable. An alternative approach employs long dsRNAs that have either been partially digested with *Escherichia coli* RNaseIII to give a mixture of short siRNAs with lengths of 18–30 nt [endoribonuclease-prepared siRNAs (esiRNAs) (20,21)] or *in vitro* digested by recombinant Dicer (22,23). The advantage of esiRNAs over synthetic siRNAs is that *in vitro* digestion of long dsRNAs results in coverage of a larger portion of the endogenous mRNA. The chances of effective sequence-specific silencing therefore increase drastically. Since this method is very cost-effective, it will likely become an invaluable tool for high-throughput screening in mammalian systems (24). Though less variable in their performance in gene silencing, esiRNAs are likely to be more susceptible to cross-silencing of homologous genes. Likewise, the current protocol for RNAi in the invertebrates *D.melanogaster* and *C.elegans* is susceptible to nonspecific silencing effects. This is due to the fact that random digestion of a large dsRNA leads to random siRNAs that will cross-react with any sequence identical to theirs. Even though there are contradictory reports on the problem of cross-silencing in RNAi (25–27), the possibility that siRNAs cross-silence genes that are identical in sequence to the intended target cannot be ruled out at this time. While the specificity of a single siRNA may be checked by a single sequence similarity search against the transcriptome or genome of the organism being studied, endoribonuclease preparation of long dsRNAs requires each potential siRNA to be analyzed for its specificity for the target, and possible cross-reactions with other genes must be excluded.

In this paper, we describe a new tool that has been specifically tailored to meet the needs of RNAi using endoribonuclease-prepared siRNAs. The program, called DEQOR, mimics esiRNAs by fragmenting the input sequence into pieces of 16–25 nt, whereby the sequence window is shifted along the input query by 1 nt at each iteration step of the algorithm. Subsequently, each *in silico* siRNA is (i) analyzed using state-of-the-art parameters for its ability to induce sequence-specific gene silencing and (ii) analyzed for its ability to cross-silence genes different from the target by performing BLAST searches against the transcriptome or genome of the organism under study. DEQOR represents the first tool that enables researchers to analyze esiRNAs, and likewise synthetic siRNAs, for their quality in terms of induction of gene silencing and for their ability to cross-silence genes other than the target itself in a high-throughput manner. A web-based user interface makes the usage and the interpretation of results retrieved by DEQOR easy and straightforward. DEQOR can be accessed freely at http://cluster-1.mpi-cbg.de/Deqor/deqor.html.

MATERIALS AND METHODS

Computational methods

BLAST searches carried out in the program DEQOR were performed using the program BLASTN from the NCBI stand-alone BLAST package (version 2.2.6) (28) with standard settings and no filtering. The web-based front-end of the program, as well as the *in silico* digestion of the input query, quality control algorithm and output parsing scripts, were written in the Python programming language. The applet for interactive selection of sequence and manipulation of penalty scores from the graphical output was programmed in Java.

DEQOR analysis of 3500 randomly selected human mRNAs

A total of 3500 mRNA clones were randomly selected from a UniGene set of 60 000 human mRNAs. Each clone was subjected to a DEQOR analysis using the following parameters: (i) cut-off E-value for target recognition 1×10^{-70}; (ii) penalty for reverse asymmetry: G/C 5′ and A/T 3′ of anti-sense strand: 7 points; A/T or G/C at both ends: 3 points; (iii) penalty for poly-G, poly-C, poly-A and poly-T: 7 points; (iv) penalty for percentage of GC-content: 1 point per 1% deviation below 20% and 1 point per 2% deviation above 50% GC; (v) penalty for perfect match to gene other than the intended target: 10 points; and (vi) penalty for match to a different gene with one mismatch: 8 points for a de-central mismatch, interpolated to 2 points for a central mismatch.

Quality parameters and cross-silenced genes were retrieved from the resulting output and analyzed further with Excel (Microsoft).

RESULTS AND DISCUSSION

DEQOR, the program

The workflow of DEQOR is shown in Figure 1. The first step in the DEQOR analysis is the identification of the source sequence. To this end, a regular BLASTN search is carried out against the selected database, using the full-length input sequence. The cut-off E-value for a positive identification is by default 1×10^{-70}. This value can, however, be adjusted by the user. The source sequence is excluded from further analysis. This step is crucial for a DEQOR analysis, since it prevents unjustified classification of cross-silencing activity against the input sequence. It should be noted that recognition of the origin does not involve the identification of potential mismatches against the target sequence. The program therefore assumes the target sequence to be identical to the input query. This can, however, be validated by the user by accessing the full-length BLAST search output from the DEQOR analysis. In the next step, the input query is fragmented *in silico* into small pieces, whereby a step-size of 1 nt is used to slide along the sequence. In this process, *in silico* siRNAs are produced which are used for further analysis. Since the size of siRNAs can vary depending on the organism or the protocol of partial digestion with RNaseIII, the user can individually select a window size between 16 and 25 nt. For each siRNA the program then (i) analyzes its silencing capability based on its base composition and (ii) identifies potential cross-silenced genes.
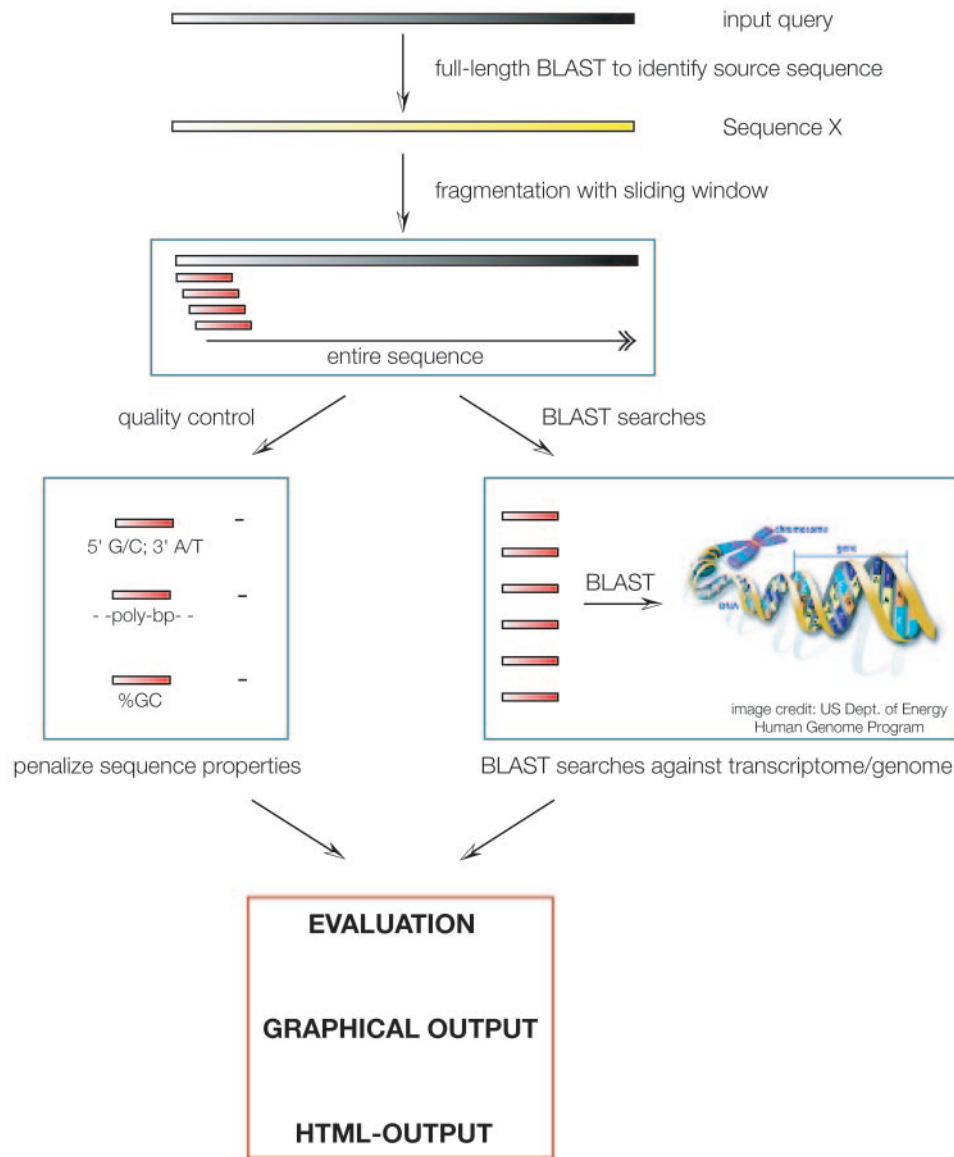
Figure 1. Workflow of a DEQOR analysis. First, the input query is submitted to a regular BLASTN search against the selected transcriptome to identify the origin of the input query. Second, the input query is digested *in silico* into small sequence pieces (between 16 and 25 nt), mimicking siRNAs. Third, each *in silico* siRNA is submitted to a BLASTN search against the selected transcriptome. Each siRNA is further analyzed for its sequence properties (see text), and disadvantageous sequence features are penalized. Finally, the siRNAs resulting from an input sequence are sorted according to their quality score and their cross-silencing capacities.

Quality control of siRNAs. Each *in silico* siRNA is subjected to quality control using state-of-the-art parameters. It was shown previously that certain sequence properties of siRNAs lead to more efficient knock-down of target genes. There are currently three major properties that are thought to be required for efficient induction of RNAi: (i) the siRNA should be asymmetric, with an A/T at its 5′ end and a G/C at the 3′ end (29,30); (ii) not more than three consecutive guanines, cytosines, adenines or thymines should occur in the siRNA sequence; and (iii) the content of guanines and cytosines (GC-content) of the siRNA sequence should be within the range of 20–50% (Frank Buchholz, unpublished data). DEQOR penalizes each siRNA according to those three criteria by a simple penalty scheme: (i) in case the siRNA has a 'reverse asymmetry', with a G/C at its 5′ end and an A/T at its 3′ end, the sequence is penalized 7 points, in case the siRNA has either an A/T or a G/C at each end, the sequence is penalized 3 points; (ii) the occurrence of more than 3 consecutive Gs, Cs, As or Ts in the sequence is penalized 7 points; (iii) finally, in case the GC-content of the sequence is below or above the specified range (20–50%), each deviation from the specified range is penalized with 1 point per 1% deviation below 20% GC-content or per 2% deviation above 50%. Default penalty settings used by DEQOR are listed in Table 1. The penalties of each siRNA are summed up to a quality score, and the siRNAs are sorted by score, with those with quality scores of zero at the top of the list and a cut-off quality score of five for good silencing capacity. The default cut-off score of five

Table 1. Default settings used by DEQOR

	Penalty
Sequence quality	
Reverse asymmetry	7
Symmetry	3
Polynucleotide stretch	7
% GC	
>50%	1 per 2% deviation
<20%	1 per 1% deviation
Cross-silencing	
Perfect match	10
One mismatch	
Central	8
De-central	2

has the consequence of selecting siRNAs with the following properties: (i) asymmetry: the siRNA will either display the correct asymmetry (A/T at 5′ end of anti-sense strand and G/C at 3′ end) or have symmetric 3′ and 5′ ends; (ii) polynucleotide stretches: the siRNA will not contain more than three consecutive As, Gs, Cs or Ts; and (iii) GC-content: the siRNA will have a GC-content between 15% and 60%.

Penalty parameters were tested in a high-throughput manner using 3500 randomly selected sequences from the human UniGene dataset by penalizing both asymmetry and polynucleotide stretches with varying penalties from 0 to 15. Several scenarios were furthermore tested for penalizing the GC-content, using default parameters for reverse asymmetry or symmetry of the siRNA and polynucleotide stretches. The results of the parameter tests are accessible as Supplementary Material (Figure S1A–D).

To examine the performance of DEQOR under the chosen standard settings, we performed a DEQOR analysis for three mammalian genes—clathrin light chain (LC), cdk2 and c-myc—all of which were previously knocked down using esiRNAs (21). According to published data, LC and cdk2 could be effectively silenced, while knock-down of c-myc was less efficient, as judged by the band intensities of the western blots. These data are reflected in the DEQOR analysis of the respective genes. Both LC and cdk2 have over 30% high-quality siRNAs and over 10% of siRNAs that meet all quality criteria; c-myc had only 15% high-quality siRNAs and 5.7% of siRNAs meeting all quality criteria (Supplementary Figure S3A–C). It should be noted that the quality parameters employed by DEQOR will be adjusted to track current knowledge on sequence criteria of siRNAs' silencing efficiency, and it should therefore not be seen as a static program.

Identification of potential cross-silencers. Each *in silico* siRNA is used as an input query for a BLASTN search against the selected nucleotide database. The identified database sequence, the percentage of identical residues between query and subject and the *E*-value of the alignment are retrieved from the BLAST output. siRNAs that show either a perfect match or a match with one or two mismatches are indicated in the DEQOR output (Figure 2C). Each cross-silencing siRNA is penalized according to the following scheme: in case of a perfect match to a gene other than the origin, a penalty of 10 points is given; in case a match with a de-central mismatch is found, the siRNA sequence is penalized
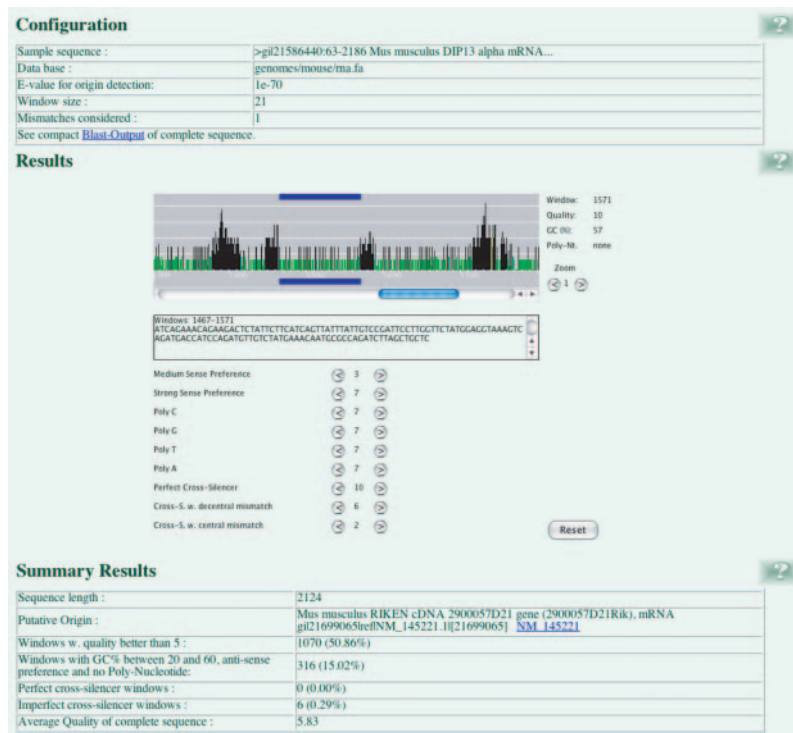
8 points; this penalty is interpolated to the center of the oligo to a penalty of 2 points. The scoring system is based on the assumption that an siRNA with a de-central match might still hybridize with the respective mRNA, but that in case of a central mismatch, hybridization is unlikely. For cross-silencing siRNAs, the accession number of the affected gene is indicated and linked to the source database (NCBI, ENSEMBL).

Owing to the high number of BLAST jobs that have to be carried out for a single DEQOR analysis, the program was parallelized on a Linux cluster (Supplementary Figure S2).

User-specific settings. Since the quality criteria for good silencing qualities of siRNAs are still in the testing phase and each laboratory might have its own preferred recipe by which siRNAs are designed, the scoring system can be individually adjusted by the user. Hence, the program offers high flexibility.

The DEQOR output. A typical output from a DEQOR analysis is shown in Figure 2. In the first section, the *Configuration* of the DEQOR analysis is displayed, followed by a short summary of the *Results* (Figure 2A). In the *Configuration window*, several parameters of the DEQOR search are listed: (i) the fasta header of the input sequence; (ii) the selected database for the BLAST search; (iii) the *E*-value for the detection of the source sequence; (iv) the window size; and (v) the number of mismatches considered for cross-silencing windows. The user is therefore reminded of the settings that were chosen for the DEQOR analysis. The *Results section* gives a graphical representation of the quality score of each siRNA along the length of the sequence, as well as a short summary and some statistical parameters of the DEQOR search. The *Graphical display* highlights siRNAs with scores higher than 5 (black), cross-silencing siRNAs (perfect match: red; match with one or more mismatches: yellow) and high-quality siRNAs in terms of gene silencing (those with a score below 5: green). By selecting a range of the graph, the user can directly access the sequence from the graphical output for further usage. At a zoom factor of 1, the position and quality parameters of the selected siRNA are shown on the right-hand side of the graph. The user can furthermore interactively manipulate the penalty parameter settings within the Java applet, resulting in reconfiguration of the quality graph. In the *Summary section*, the program gives the following information on the results of the DEQOR analysis: (i) the length of the input query; (ii) the putative origin of the input query (linked to the NCBI/ ENSEMBL sequence entry); (iii) the number and relative number of siRNAs with a quality score better than 5; (iv) the number and relative number of siRNAs that meet all quality criteria according to the input settings; (v) the number and relative number of siRNAs that show perfect cross-silencing activity; (vi) siRNAs that have cross-silencing capabilities with a single mismatch; and finally (vii) the average gene-silencing quality of the entire input query. This summary of quality parameters is especially helpful for esiRNA-based gene silencing, since the overall quality of the input sequence might indicate the probable success rate of an RNA interference experiment. If the query has a large number of high-quality siRNAs, the success rate of gene silencing will increase. If the majority of siRNAs produced by a sequence are low quality, the researcher might consider using a different portion of the

Figure 2. Typical output of a DEQOR analysis. The output window is divided into four major parts. (**A**) A summary section displays the chosen configuration of the DEQOR run, an interactive graphical display of the quality scores of siRNAs along the input sequence, the putative origin of the input sequence and its length and finally a statistical analysis of the overall quality and cross-silencing capacities. (**B**) The 10 top quality windows are listed, showing their position in the sequence ('Window'), the sequence of the siRNA ('Sequence', with anti-sense strand in boldface), the number of perfect matches ('Perf. M.'), matches with one mismatch ('w/1 Mism.', the percentage of GC–content ('GC%'), the asymmetric features of the siRNA ('ASP'; anti-sense preference, with the options 'yes', 'no' and 'sym.'), the occurrence of more than three consecutive identical nucleotides in the siRNA sequence ('Poly') and finally the quality score of the siRNA ('Quality'). (**C**) Windows that are potential cross-silencers are shown. The columns are identical to (**B**). The accession numbers (NCBI/ENSEMBL) of genes that are targets for cross-silencing activities are listed and linked to their source database at the NCBI/ENSEMBL. Mouse Dip13 α (NM_145221) was used for this DEQOR analysis.

gene. Finally, cross-silencing activity of a mixture of esiRNAs might be reduced to a minimum by excluding regions with cross-silencing activity from the preparation of esiRNAs.

In the next section of the output, the 10 *Top quality windows* are shown (with a cut-off score of 5 used to indicate good silencing potential), giving information on the location of the siRNA and its sequence, potential cross-silenced genes, the GC-content, whether the siRNA lacks asymmetry or contains reverse asymmetry and whether it contains a polynucleotide stretch, and finally the score of the siRNA (Figure 2B). Next, all *Cross-silencers* of siRNAs are listed, including the accession number and a link to the cross-silenced gene. All information shown in the top quality window section is also given for cross-silencing siRNAs (Figure 2C). Finally, a summary of the full-length BLAST search with the input query is shown, giving information about identified genes, percentage of identical residues and the *E*-value for each hit (data not shown). The full-length BLAST output including alignments is furthermore linked in this section.

Biological implication of high-throughput DEQOR searches

In order to estimate average quality parameters and cross-silencing capacities for a large dataset of mRNAs, we performed a DEQOR analysis for approximately 3500 human mRNA fragments that were randomly selected from a total of 60 000 sequenced clones. DEQOR analysis was carried out using default parameters. From the resulting output we retrieved several parameters: (i) the percentage of siRNAs per gene that had a quality score below 5; (ii) the percentage of siRNAs that met all quality criteria according to standard settings; (iii) the percentage of perfect cross-silencers per gene; (iv) the percentage of cross-silencers with one mismatch per gene; (v) the average quality of the complete input sequence; (vi) the average quality of the complete input sequence summarized with cross-silencing penalties; and finally (vii) the length of the input sequence. The statistics of the DEQOR analysis are shown in Figure 3. The percentage of siRNAs per gene that had a quality score below 5 showed approximately a normal distribution, with a peak between 40 and 50% of siRNAs per gene (with close to 30% of genes). On the other hand, only a fraction of siRNAs per gene strictly met all quality criteria (siRNA sequence shows correct asymmetry; no polynucleotide stretch within the sequence; GC-content between 15 and 60%). Of the genes, 60% contained only 10% of siRNAs that met these criteria, and 40% of genes contained 20% of siRNAs meeting this high standard (Figure 3A). We were also interested in the average quality of the selected sequences in our dataset. To this end, we calculated the average quality score of each complete input query (i) without cross-silencer penalties and (ii) summarized with cross-silencer penalties (Figure 3B). Only 4% of all input sequences had an average quality score below 5, 80% of genes displayed a quality score between 5 and 10, and 15% had an average quality score above 10. The average quality scores did not differ greatly when cross-silencer penalties were considered (3.5% of genes had a quality score below 5, 75% between 5 and 10, 19.5% between 10 and 20 and, finally, 2.6% above 20). These data suggest that essentially all genes in our dataset should be sufficiently effective in esiRNA-based RNA

interference. Finally, we were interested in how many cross-silencers with a quality score better than 5 were present in our dataset. In essence, we assumed that only those siRNAs that showed good silencing qualities would be potent cross-silencers. In addition, we assumed that, due to the high number of different siRNAs, only those genes that contain a substantial amount of cross-silencing siRNAs would be candidates for nonspecific silencing. To this end, we calculated the percentage of cross-silencing siRNAs per gene that had a quality score better than 5 (Figure 3C). Of all genes 91.5% did not contain a single cross-silencer with a quality score below 5, and only 3.4% of genes contained more than 10 high-quality cross-silencers. These results are especially encouraging for high-throughput esiRNA-based gene silencing, since <10% of all genes showed any cross-silencing activity. To discover whether high numbers of cross-silencing siRNAs were due to binding of oligos to the 3′- and 5′-untranslated regions (UTRs) of the gene, we analyzed further the high-quality cross-silencers for their binding position in the cross-silenced gene (data not shown). In 24.4% of cases, cross-silencing siRNAs were predicted to recognize a sequence in both the coding sequence and one of the UTRs. In 50.8% of all cases, only a part of the coding sequence was hit, and 24.8% of the genes showed reaction with one of the UTRs only.

When analyzing the data, we realized that the recognition of the target is one of the inherent problems in running a DEQOR analysis, since the cut-off value of 1×10^{-70} is in some cases not met by the input sequence, due to either low-complexity regions within the input query or, more often, the shortness of the query. If DEQOR is used for analysis of single genes, this will not cause a problem. If, however, the program is used for high-throughput screening of a large dataset, the output has to be analyzed carefully.

Applications for DEQOR

There are two basic applications for DEQOR. First, DEQOR provides a fast, efficient and easy-to-use tool for analyzing primer qualities of siRNAs for a gene. In this respect, the program may also be used for the design of siRNAs for chemical synthesis. The researcher might, for instance, display only the 10 best windows in terms of gene-silencing capacities. The second possible application is to use DEQOR for the analysis of the gene-silencing efficiency and cross-silencing capacities of genes to be used for esiRNA-based gene silencing, since it evaluates each potential siRNA resulting from a sequence for its silencing capabilities, as well as for its cross-silencing activity by performing BLAST searches against the transcriptome or genome. After running a DEQOR analysis with a gene of choice, the researcher might select only those fragments of a sequence that contain a high number of efficient siRNAs and do not contain cross-silencing windows. At the moment, vertebrate (human, mouse and rat), as well as invertebrate (*D. melanogaster* and *C.elegans*), transcriptomes and the respective genomes are available for searching. The UniGene datasets of the plants *Arabidopsis thaliana*, *Oryza sativa* and *Zea mays* are furthermore supported by the DEQOR server. When analyzing invertebrate siRNAs, it is worth mentioning that while RNaseIII from *E.coli* cuts dsRNA randomly, this is not the case with Dicer (31–33). Dicer-like RNases preferentially digest long, double-stranded RNAs
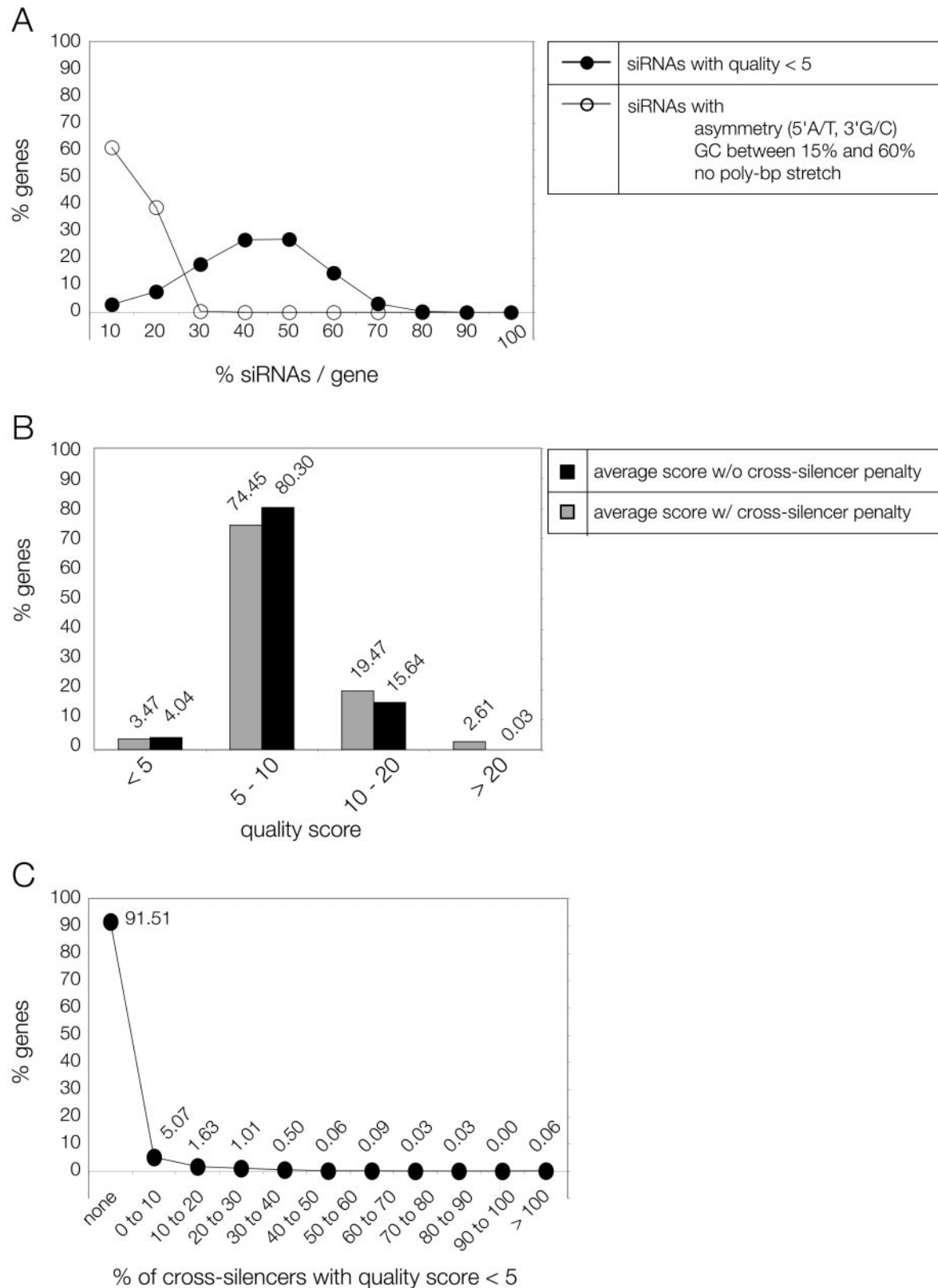
Figure 3. Statistical analysis of esiRNAs from 3500 randomly selected human genes. (**A**) The % of siRNAs per gene that have a quality score better than 5 (closed circles) and the % siRNAs that meet all standard quality criteria (open circles) were plotted against the % genes. (**B**) The average quality scores of each input sequence without (dark gray) and with (light gray) cross-silencing capabilities were related to the % genes in the dataset. Quality scores were divided into four categories: (i) below 5; (ii) between 5 and 10; (iii) between 10 and 20; and (iv) above 20. (**C**) The % high-quality cross-silencers is plotted against the % genes in the dataset. Only 8.49% of genes have one or more cross-silencers with a quality score below 5 (see text for further discussion).

from their termini, resulting in pieces of ∼22 nt in length. Since, however, it is unclear whether this type of digestion is not at least partially random, we decided to use the same type of *in silico* digestion for these two organisms, using a step size of 1 nt to move along the sequence.

In conclusion, DEQOR should be a useful tool for experimentalists for analyzing the quality and specificity of siRNAs and esiRNAs.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

ACKNOWLEDGEMENTS

REFERENCES

1. Hannon,G.J. (2002) RNA interference. *Nature*, **418**, 244–251.
2. Denli,A.M. and Hannon,G.J. (2003) RNAi: an ever-growing puzzle. *Trends Biochem Sci.*, **28**, 196–201.
3. Cogoni,C. and Macino,G. (2000) Post-transcriptional gene silencing across kingdoms. *Curr. Opin. Genet. Dev.*, **10**, 638–643.
4. Sharp,P.A. and Zamore,P.D. (2000) Molecular biology. RNA interference. *Science*, **287**, 2431–2433.
5. Sharp,P.A. (2001) RNA interference—2001. *Genes Dev.*, **15**, 485–490.
6. Plasterk,R.H. and Ketting,R.F. (2000) The silence of the genes. *Curr. Opin. Genet. Dev.*, **10**, 562–567.
7. Zamore,P.D. (2001) RNA interference: listening to the sound of silence. *Nature Struct. Biol.*, **8**, 746–750.
8. Boutros,M., Kiger,A.A., Armknecht,S., Kerr,K., Hild,M., Koch,B., Haas,S.A., Consortium,H.F., Paro,R. and Perrimon,N. (2004) Genome-wide RNAi analysis of growth and viability in *Drosophila* cells. *Science*, **303**, 832–835.
9. Piano,F., Schetter,A.J., Morton,D.G., Gunsalus,K.C., Reinke,V., Kim,S.K. and Kemphues,K.J. (2002) Gene clustering based on RNAi phenotypes of ovary-enriched genes in *C. elegans*. *Curr. Biol.*, **12**, 1959–1964.
10. Pothof,J., van Haaften,G., Thijssen,K., Kamath,R.S., Fraser,A.G., Ahringer,J., Plasterk,R.H. and Tijsterman,M. (2003) Identification of genes that protect the *C. elegans* genome against mutations by genome-wide RNAi. *Genes Dev.*, **17**, 443–448.
11. Vastenhouw,N.L., Fischer,S.E., Robert,V.J., Thijssen,K.L., Fraser,A.G., Kamath,R.S., Ahringer,J. and Plasterk,R.H. (2003) A genome-wide screen identifies 27 genes involved in transposon silencing in *C. elegans*. *Curr. Biol.*, **13**, 1311–1316.
12. Parrish,J.Z. and Xue,D. (2003) Functional genomic analysis of apoptotic DNA degradation in *C. elegans*. *Mol. Cell.*, **11**, 987–996.
13. Kamath,R.S. and Ahringer,J. (2003) Genome-wide RNAi screening in *Caenorhabditis elegans*. *Methods*, **30**, 313–321.
14. Herold,A., Teixeira,L. and Izaurralde,E. (2003) Genome-wide analysis of nuclear mRNA export pathways in *Drosophila*. *EMBO J.*, **22**, 2472–2483.
15. Ashrafi,K., Chang,F.Y., Watts,J.L., Fraser,A.G., Kamath,R.S., Ahringer,J. and Ruvkun,G. (2003) Genome-wide RNAi analysis of *Caenorhabditis elegans* fat regulatory genes. *Nature*, **421**, 268–272.
16. Williams,B.R. (1999) PKR; a sentinel kinase for cellular stress. *Oncogene*, **18**, 6112–6120.
17. Stark,G.R., Kerr,I.M., Williams,B.R., Silverman,R.H. and Schreiber,R.D. (1998) How cells respond to interferons. *Annu. Rev. Biochem.*, **67**, 227–264.
18. Elbashir,S.M., Harborth,J., Lendeckel,W., Yalcin,A., Weber,K. and Tuschl,T. (2001) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature*, **411**, 494–498.
19. Holen,T., Amarzguioui,M., Wiiger,M.T., Babaie,E. and Prydz,H. (2002) Positional effects of short interfering RNAs targeting the human coagulation trigger Tissue Factor. *Nucleic Acids Res.*, **30**, 1757–1766.
20. Calegari,F., Haubensak,W., Yang,D., Huttner,W.B. and Buchholz,F. (2002) Tissue-specific RNA interference in postimplantation mouse embryos with endoribonuclease-prepared short interfering RNA. *Proc. Natl Acad. Sci., USA*, **99**, 14236–14240.
21. Yang,D., Buchholz,F., Huang,Z., Goga,A., Chen,C.Y., Brodsky,F.M. and Bishop,J.M. (2002) Short RNA duplexes produced by hydrolysis with *Escherichia coli* RNase III mediate effective RNA interference in mammalian cells. *Proc. Natl Acad. Sci., USA*, **99**, 9942–9947.
22. Myers,J.W., Jones,J.T., Meyer,T. and Ferrell,J.E.,Jr (2003) Recombinant Dicer efficiently converts large dsRNAs into siRNAs suitable for gene silencing. *Nature Biotechnol.*, **21**, 324–328.
23. Kawasaki,H., Suyama,E., Iyo,M. and Taira,K. (2003) siRNAs generated by recombinant human Dicer induce specific and significant but target site-independent gene silencing in human cells. *Nucleic Acids Res.*, **31**, 981–987.
24. Kittler,R. and Buchholz,F. (2003) RNA interference: gene silencing in the fast lane. *Semin. Cancer Biol.*, **13**, 259–265.
25. Semizarov,D., Frost,L., Sarthy,A., Kroeger,P., Halbert,D.N. and Fesik,S.W. (2003) Specificity of short interfering RNA determined through gene expression signatures. *Proc. Natl Acad. Sci., USA*, **100**, 6347–6352.
26. Chi,J.T., Chang,H.Y., Wang,N.N., Chang,D.S., Dunphy,N. and Brown,P.O. (2003) Genomewide view of gene silencing by small interfering RNAs. *Proc. Natl Acad Sci., USA*, **100**, 6343–6346.
27. Jackson,A.L., Bartz,S.R., Schelter,J., Kobayashi,S.V., Burchard,J., Mao,M., Li,B., Cavet,G. and Linsley,P.S. (2003) Expression profiling reveals off-target gene regulation by RNAi. *Nature Biotechnol.*, **21**, 635–637.
28. Altschul,S.F., Madden,T.L., Schaffer,A.A., Zhang,J., Zhang,Z., Miller,W. and Lipman,D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.
29. Khvorova,A., Reynolds,A. and Jayasena,S.D. (2003) Functional siRNAs and miRNAs exhibit strand bias. *Cell*, **115**, 505.
30. Schwarz,D.S., Hutvagner,G., Du,T., Xu,Z., Aronin,N. and Zamore,P.D. (2003) Asymmetry in the assembly of the RNAi enzyme complex. *Cell*, **115**, 199–208.
31. Ketting,R.F., Fischer,S.E., Bernstein,E., Sijen,T., Hannon,G.J. and Plasterk,R.H. (2001) Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev.*, **15**, 2654–2659.
32. Zamore,P.D., Tuschl,T., Sharp,P.A. and Bartel,D.P. (2000) RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell*, **101**, 25–33.
33. Zhang,H., Kolb,F.A., Brondani,V., Billy,E. and Filipowicz,W. (2002) Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. *EMBO J.*, **21**, 5875–5885.